

# There Are (super)Giants in the Sky: Searching for Misidentified Massive Stars in Algorithmically-Selected Quasar Catalogs

Trevor Z. Dorn-Wallenstein and Emily Levesque

Astronomy Department, University of Washington,  
 Physics and Astronomy Building, 3910 15th Ave NE, Seattle, WA 98105, USA  
 email: tzdw@uw.edu, emsque@uw.edu

## Abstract.

Thanks to incredible advances in instrumentation, surveys like the Sloan Digital Sky Survey have been able to find and catalog billions of objects, ranging from local M dwarfs to distant quasars. Machine learning algorithms have greatly aided in the effort to classify these objects; however, there are regimes where these algorithms fail, where interesting oddities may be found. We present here an X-ray bright quasar misidentified as a red supergiant/X-ray binary, and a subsequent search of the SDSS quasar catalog for X-ray bright stars misidentified as quasars.

**Keywords.** astronomical data bases: surveys, stars: supergiants, X-rays: binaries

## 1. Introduction/Overview

### 1.1. Red Supergiant X-ray Binaries

Over the past decade, many exotic close binary systems with supergiant components have been discovered. Systems like NGC 300 X-1 — a Wolf-Rayet/black hole X-ray binary (Crowther *et al.* 2010) — and SN2010da — a sgB[e]/neutron star X-ray binary (Villar *et al.* 2016) — are two such examples of a coupling between a massive star in a short-lived evolutionary phase and a compact stellar remnant. Interestingly, no X-ray binaries with confirmed red supergiant (RSG) counterparts have been discovered (RSGs have been proposed as the candidate donor star for a few Ultraluminous X-ray Sources, see Heida *et al.* 2016). This may be partially explained by the rarity of RSGs; however, though rare, RSGs are both longer-lived and more common (due to the smaller — 10 - 25  $M_{\odot}$  — initial masses of their zero age main sequence progenitors) than most other evolved massive stars.

RSG X-ray binaries, if they exist, offer a view into an interesting edge case of accretion; their extended envelopes and strong winds ( $\dot{M} \sim 10^{-4} M_{\odot} \text{ yr}^{-1}$ , van Loon *et al.* 2005) could allow for accretion from both the wind and Roche-Lobe Overflow in an environment continually enriched with dust produced by the RSG. RSG X-ray binaries are also the immediate progenitors of Thorne-Żytkow Objects — stars with embedded neutron star cores (Thorne & Żytkow 1975) — assuming the neutron star plunges into the RSG as it expands (Taam *et al.* 1978).

### 1.2. J0045+41

To search for RSG X-ray binaries, we used the photometry of the Local Group Galaxy Survey (LGGS, Massey & Olsen 2003, Massey *et al.* 2006, 2007), which covers M31, M33, the Magellanic Clouds and 7 dwarf galaxies in the Local Group. Following Massey (1998) to find RSGs among the nearly-identical foreground dwarfs, we cross-referenced the po-

sitions of the LGGS RSGs with the *Chandra* Source Catalog (CSC, Evans *et al.* 2010), and found one RSG coincident with an X-ray source.

LGGS J004527.30+413254.3 (J0045+41 hereafter) is a bright ( $V \approx 19.9$ ) object of previously-unknown nature in the disk of M31. Vilardell *et al.* (2006) classify J0045+41 as an eclipsing binary with a period of  $\sim 76$  days. J0045+41 was also observed with the Palomar Transient Factory (PTF); the  $g$ -band lightcurve shows evidence for a  $\sim 650$  day period. On the other hand Kim *et al.* (2007) identify J0045+41 as a globular cluster, and it has been included in catalogs of M31 globular clusters as recently as 2014 (Wang *et al.* 2014). The LGGS photometry is consistent with the color and brightness of a RSG. Indeed, following Levesque *et al.* (2006), we found that, as an RSG, J0045+41 would have an effective temperature of  $\sim 3500$  K and bolometric magnitude of  $-6.67$ , consistent with a  $12\text{--}15 M_{\odot}$  RSG. However, a complete SED fit to photometry from the Panchromatic Hubble Andromeda Treasury (PHAT, Dalcanton *et al.* 2012) using the Bayesian Extinction and Stellar Tool (BEAST, Gordon *et al.* 2016) yields an unphysical result of  $300 M_{\odot}$ ,  $10^5$  K star, extinguished by  $A_V \sim 4$  magnitudes. Furthermore, the object appears extended in the PHAT images (though its radial profile appears similar to that of other nearby stars).

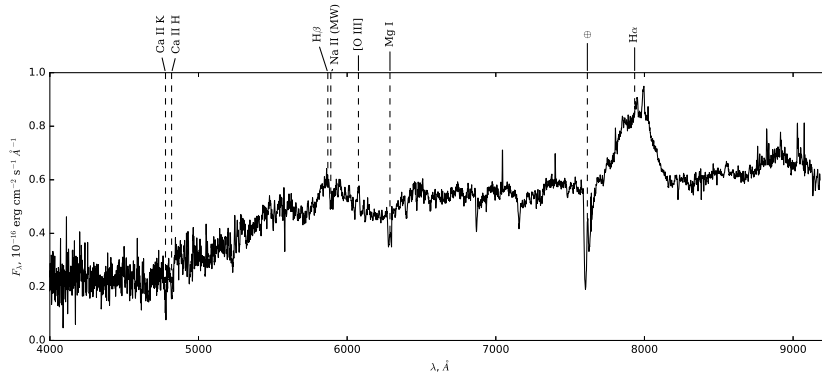
J0045+41 is separated by  $\sim 1.18''$  from an X-ray source. The source, CXO J004527.3+413255, is bright ( $F_X = 1.98 \times 10^{-13}$  erg s $^{-1}$  cm $^{-2}$ ) and hard; fitting a spectrum obtained by Williams *et al.* (in prep) yields a power law with  $\Gamma \sim 1.5$ . The best-fit neutral hydrogen column density is  $1.7 \times 10^{21}$  cm $^{-2}$ , which corresponds to  $A_V \sim 1$ .

### 1.2.1. Observations and Data Reduction

The apparent periodicity of J0045+41 and its apparent association with a hard and unabsorbed X-ray source prompted us to obtain follow-up spectroscopic observations to determine the true nature of this object. We obtained a longslit spectrum of J0045+41 using the Gemini Multi-Object Spectrograph (GMOS) on Gemini North. Four 875 second exposures were taken 2016 July 5 using the B600 grating centered on  $5000 \text{ \AA}$ , and four 600 second exposures were taken 2016 July 9 using the R400 grating centered on  $7000 \text{ \AA}$ , with a blocking filter to remove  $2^{nd}$ -order diffraction. Due to the gaps between GMOS's three CCDs, two of each set of exposures were offset by  $+50 \text{ \AA}$ . The data were reduced using the standard `gemini IRAF` package. The final reduced spectrum has continuous signal from  $\sim 4000$  to  $\sim 9100 \text{ \AA}$  at a resolution of  $R \sim 1688$  (blue)/1918 (red).

### 1.2.2. Spectrum and Redshift Determination

The spectrum (Figure 1) shows that J0045+41 is a quasar at  $z \approx 0.21$  (measured with  $H\alpha$ ,  $H\beta$ , [OIII] $\lambda 5007$  and Ca II H and K). While a false positive, this quasar is quite interesting on its own. The (low-significance) detections of periodicity on short timescales by multiple sources are difficult to explain. Furthermore, mistaking a blue quasar for a red star would imply a high reddening along the line of sight, which is consistent with a sightline through M31; however, the low H column density implied by the X-ray counterpart's spectrum and our inability to achieve a satisfactory fit to the optical spectrum by reddening the (redshifted) quasar template spectrum from Vanden Berk *et al.* (2001) indicate that J0045+41 belongs to a small and intriguing class of intrinsically red quasars, observed through a low-extinction region of the ISM in M31 (Richards *et al.* 2003).



**Figure 1.** GMOS spectrum of J0045+41 with all identified lines labeled. The Na II feature is intrinsic to the Milky Way/M31, and the  $\oplus$  line is telluric.

## 2. Stars in the SDSS Quasar Catalog

If a quasar can be misidentified as a RSG, are there red stars — especially X-ray bright stars — in already-existing quasar catalogs? To answer this question, we turned to the quasar catalog of the Sloan Digital Sky Survey (SDSS, York *et al.* 2000).

### 2.1. Sample Selection

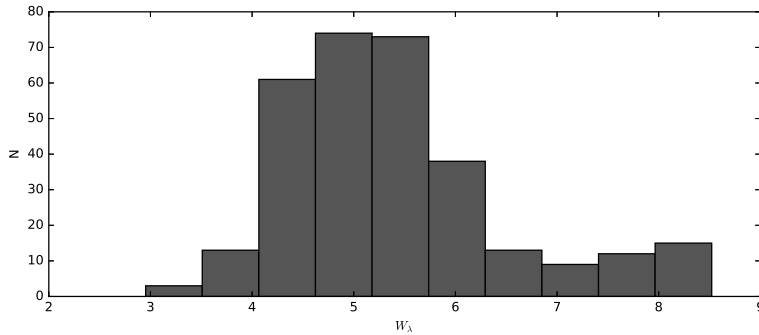
We selected all SDSS objects automatically tagged as quasars that were within 0.2 magnitudes of J0045+41 in  $g - r$  vs.  $r - i$  vs.  $i - z$  color-space — J0045+41 is too faint in  $u$  to utilize  $u - g$  — and ignored any warning flags to avoid throwing out interesting objects that were not easily identified by the SDSS algorithm. 1098 objects in this sample had associated spectroscopic observations. Interestingly, many of the spectroscopically-determined redshifts were unbelievably small or even negative, implying that these objects are in a regime of color-space where classification algorithms may fail. Indeed, on visual inspection of these spectra, many of them are stellar.

### 2.2. Stars

We used `emcee`, a Python implementation of Markov Chain Monte Carlo (MCMC) by Foreman-Mackey *et al.* (2013), to fit Gaussian profiles to the Ca II triplet ( $\lambda = 8498, 8542, \text{ and } 8662 \text{ \AA}$ ), which we use to identify stars. The posterior distributions of the parameters allow us to determine if the triplet is well fit, and estimate errors for each parameter. Because the relative centroids and strengths of the lines are fixed, a good fit guarantees the lines are actually detected, while simultaneously measuring — with accurate errors — the radial velocity and equivalent width ( $W_\lambda$ ) of the triplet. After a follow-up inspection by eye of spectra that were noisy or missing data, we find 344 confirmed cool stars, representing  $\sim 31\%$  of the total sample. Figure 2 shows the distribution of  $W_\lambda$  for  $W_\lambda/\sigma_{W_\lambda} > 1$ . We follow Jennings & Levesque (2016) to estimate luminosity from  $W_\lambda$ , and find that most stars are dwarfs ( $W_\lambda \lesssim 6.5 \text{ \AA}$ ) but  $\sim 40$  stars have larger equivalent widths indicating they are likely giants or supergiants (the relationship is dependent on effective temperature and metallicity, so these labels are approximate).

## 3. Discussion and Future Work

This result demonstrates that, when looking for rare objects like RSG X-ray binaries, it is important to look in unlikely places; e.g., a red and X-ray bright star may be confused



**Figure 2.** Distribution of measured equivalent widths for confirmed stars.

for a quasar if the classification algorithm mistakes the continuum between the TiO bands in a RSG spectrum for an emission line. The fact that some of the color-space containing M-dwarfs — by far the most common type of star — is a regime where classification algorithms fail underlines the importance of improving on these algorithms until they perform as well as the human eye. Indeed, many of these stars were previously identified (see West *et al.* 2011), but are still listed as quasars on the SDSS online data portal.

Future work will focus on improving our star-finding algorithm to use alternate spectral features when the Ca II triplet is missing or obscured by noise, and on finding which areas of color-space contain significant numbers of these misidentified stars, with the goal of finding RSG X-ray binaries as well as improving our knowledge of where classification algorithms fail.

## References

- Crowther, P. A., Barnard, R., Carpano, S., *et al.* 2010, *MNRAS*, 403, L41  
Dalcanton, J. J., Williams, B. F., Lang, D., *et al.* 2012, *ApJS*, 200, 18  
Davidsen, A., Malina, R., & Bowyer, S. 1977, *ApJ*, 211, 866  
Evans, I. N., Primini, F. A., Glotfelty, K. J., *et al.* 2010, *ApJS*, 189, 37-82  
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306  
Gordon, K. D., Fouesneau, M., Arab, H., *et al.* 2016, *ApJ*, 826, 104  
Heida, M., Jonker, P. G., Torres, M. A. P., *et al.* 2016, *MNRAS*, 459, 771  
Jennings, J., & Levesque, E. M. 2016, *ApJ*, 821, 131  
Kim, S. C., Lee, M. G., Geisler, D., *et al.* 2007, *AJ*, 134, 706  
Levesque, E. M., Massey, P., Olsen, K. A. G., *et al.* 2006, *ApJ*, 645, 1102  
Massey, P. 1998, *ApJ*, 501, 153  
Massey, P., & Olsen, K. A. G. 2003, *AJ*, 126, 2867  
Massey, P., Olsen, K. A. G., Hodge, P. W., *et al.* 2006, *AJ*, 131, 2478  
Massey, P., Olsen, K. A. G., Hodge, P. W., *et al.* 2007, *AJ*, 133, 2393  
Richards, G. T., Hall, P. B., Vanden Berk, D. E., *et al.* 2003, *AJ*, 126, 1131  
Taam, R. E., Bodenheimer, P., & Ostriker, J. P. 1978, *ApJ*, 222, 269  
Thorne, K. S., & Żytkow, A. N. 1975, *ApJL*, 199, L19  
Vanden Berk, D. E., Richards, G. T., Bauer, A., *et al.* 2001, *AJ*, 122, 549  
van Loon, J. T., Cioni, M.-R. L., Zijlstra, A. A., & Loup, C. 2005, *A&A*, 438, 273  
Vilardell, F., Ribas, I., & Jordi, C. 2006, *A&A*, 459, 321  
Villar, V. A., Berger, E., Chornock, R., *et al.* 2016, *ApJ*, 830, 11  
Wang, S., Ma, J., Wu, Z., & Zhou, X. 2014, *AJ*, 148, 4  
West, A. A., Morgan, D. P., Bochanski, J. J., *et al.* 2011, *AJ*, 141, 97  
York, D. G., Adelman, J., Anderson, J. E., Jr., *et al.* 2000, *AJ*, 120, 1579